

CLAIMS

- Sub
Att
1. A method of retraining a trainable data classifier comprising the steps of:
- 5 providing a first item of training data;
- 10 comparing the first item of training data with a second item of training data already used to train the data classifier;
- 15 calculating a measure of conflict between the first and second items of training data;
- 20 using the first item of training data to retrain the data classifier responsive to the measure of conflict.
- 25 2. A method according to claim 1 wherein the step of using the first item of training data is responsive to a predetermined conflict threshold value.
- 30 3. A method according to claim 2 wherein the threshold value is non-zero.
- 35 4. A method according to claim 1 wherein the measure of conflict comprises a geometric difference between the first and second items of training data.
5. A method according to claim 4 wherein the geometric difference comprises a Euclidean distance.
- 30 6. A method according to claim 1 wherein the measure of conflict comprises an association coefficient of the first and second items of training data.

7. A method according to claim 6 wherein the association coefficient is a Jaccard's coefficient.

8. A method according to claim 7 wherein the measure of conflict is derived from both a Euclidean distance between and a Jaccard's coefficient of the first and second items of training data.

9. A method according to claim 8 wherein the measure of conflict is derived from a Euclidean distance and a Jaccard's coefficient composed in an exponential relationship with respect to each other.

10. A method according to claim 8 wherein the measure of conflict is derived from a function of a Euclidean distance multiplied by an exponent of a function of the Jaccard's coefficient.

11. A method according to claim 1 wherein the data classifier comprises a neural network.

12. A method according to claim 1 wherein the training data comprises telecommunications network data.

13. A method according to claim 1 wherein the training data comprises telecommunications call detail record data.

14. A method of training a trainable data classifier comprising the steps of:

providing a plurality of items of training data;

comparing a first of the items of training

data with a second of the items of training
data;

5 calculating a measure of conflict between
the first and second items of training data;

10 using one of the first and second items of
training data to retrain the data classifier
responsive to the measure of conflict.

15. A apparatus for retraining a trainable
data classifier and comprising:

15 an input port for receiving a first item of
training data;

20 a comparator arranged to compare the first
item of training data with a second item of
training data already used to train the data
classifier;

25 a calculator for calculating a measure of
conflict between the first and second items of
training data; and

30 an output port arranged to output the first
item of training data to the data classifier
responsive to the measure of conflict.

30 16. A anomaly detection system comprising
apparatus according to claim 15.

35 17. A telecommunications data anomaly
detection system comprising apparatus according to
claim 15.

18. A telecommunications fraud detection

system comprising apparatus according to claim 15.

19. An account fraud detection system comprising apparatus according to claim 15.

5 20. An apparatus for restraining a trainable
data classifier comprising:

an input port for receiving a plurality of items of training data;

a comparator arranged to compare a first of the items of training data with a second of the items of training data;

15 a calculator for calculating a measure of conflict between the first and second items of training data;

20 an output port arranged to output the first item of training data to the data classifier responsive to the measure of conflict.

25 21. A program for a computer on a machine
readable medium arranged to perform the steps of:

receiving a first item of training data;

30 comparing the first item of training data
with a second item of training data already used
to train the data classifier;

calculating a measure of conflict between
the first and second items of training data;

35 using the first item of training data to
retrain the data classifier responsive to the

measure of conflict.

22. A program for a computer on a machine readable medium arranged to perform the steps of:

5 receiving a plurality of items of training data;

10 comparing a first of the items of training data with a second of the items of training data;

15 calculating a measure of conflict between the first and second items of training data; and

using one of the first and second items of training data to retrain the data classifier responsive to the measure of conflict.

Out
cont.

0
1
2
3
4
5
6
7
8
9